ABSTRACT

There are provided a multi-point conference system and a multi-point conference device capable of using an SIP terminal as a conference terminal and clearly and smoothly performing speaker switching without requesting a conference terminal to transmit an intra frame when a speaker is detected. The multi-point connection device includes an image processing unit for accumulating in a memory (13), image data from the conference terminal participating in the conference, and decoding the image data on the speaker accumulated in the memory (13) and re-encoding the decoded image upon detection of a speaker in a medium processing unit (15). Upon switching of a speaker, the image processing unit re-encodes the first image as an intra frame and transmits the image data to conference terminals other than that of the speaker, so that the conference terminals other than that of the speaker can perform decoding from the intra frame, thereby clearly and smoothly performing the speaker switching.